

REPORT DOCUMENTATION PAGE			Form Approved OMB NO. 0704-0188		
<p>The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA, 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.</p> <p>PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.</p>					
1. REPORT DATE (DD-MM-YYYY)		2. REPORT TYPE		3. DATES COVERED (From - To)	
		Technical Report		-	
4. TITLE AND SUBTITLE Final Report on "Survey of Quantification and Distance FunctionsUsed for Internet-based Weak-link Sociological Phenomena"			5a. CONTRACT NUMBER		
			W911NF-11-1-0322		
			5b. GRANT NUMBER		
6. AUTHORS Ming-Jun Lai, Dawn Robinson			5c. PROGRAM ELEMENT NUMBER		
			611102		
			5d. PROJECT NUMBER		
			5e. TASK NUMBER		
			5f. WORK UNIT NUMBER		
7. PERFORMING ORGANIZATION NAMES AND ADDRESSES University of Georgia Research Foundation, Inc. Sponsored Research Office University of Georgia Research Foundation Inc Athens, GA 30602 -			8. PERFORMING ORGANIZATION REPORT NUMBER		
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U.S. Army Research Office P.O. Box 12211 Research Triangle Park, NC 27709-2211			10. SPONSOR/MONITOR'S ACRONYM(S) ARO		
			11. SPONSOR/MONITOR'S REPORT NUMBER(S) 60673-MA.1		
12. DISTRIBUTION AVAILABILITY STATEMENT Approved for public release; distribution is unlimited.					
13. SUPPLEMENTARY NOTES The views, opinions and/or findings contained in this report are those of the author(s) and should not contrued as an official Department of the Army position, policy or decision, unless so designated by other documentation.					
14. ABSTRACT We sumarize our research activities and results in this manuscript. It eneds with some conclusions.					
15. SUBJECT TERMS research Activities, research results, conclusion					
16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	15. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			Ming-Jun Lai
UU	UU	UU	UU		19b. TELEPHONE NUMBER
					706-542-2065

Report Title

Final Report on "Survey of Quantification and Distance FunctionsUsed for Internet-based Weak-link Sociological Phenomena"

ABSTRACT

We summarize our research activities and results in this manuscript. It ends with some conclusions.

Final Report on "Survey of Quantification and Distance Functions Used for Internet-based Weak-link Sociological Phenomena" *

Ming-Jun Lai [†]

Dawn Robinson[‡]

January 16, 2013

1 Research Activities:

Dr. Ming-Jun Lai, the PI and Dr. Dawn Robinson, the co-PI had a few meetings in the Lai's office and one meeting at the Robinson's office and discussed how to proceed the research during the Aug. and Sept., 2011. They had several emails communications before and after these meeting. Also, later in March 2012, they had another meeting on the distribution of the grant money, one summer month salary for the PI and one summer month salary for the co-PI. The PI used all the fund allocated for travel, a part of the travel fund was used for Mr. George Slavov, one of the PI's graduate students for his summer salary and a part of the travel fund was used for Mr. Leopold Matamba-Messi's travel to UCLA for a conference. Mr. Leopold Matamba Messi is another graduate students of the PI. Mr. Matamba-Messi graduated in the Aug. 2012 and found a post-doc position at the Mathematical Biology Institute at Ohio State University. Since Aug., 2012, the PI has emailed to the co-PI about the final report several times. The PI and co-PI met again on Dec. 5, 2012 during a workshop and discussed how to write this report. They have met again in the Lai's office on Jan. 16, 2013 for submission of the final report.

2 Research Results:

The PI studied all mathematical literature he can find related to the Google search engine, Google matrix, PageRank as well as the Yahoo search engine and a classic SearchKing HIST algorithm. The co-PI immersed herself in the sociology literature for the relevant studies on social network, strong and weak social ties and etc.. which may be useful for the study of search algorithms. Our findings can be summarized as follows.

- 1) A review of the related mathematical literature is given in section 2.2 of Appendix I.
- 2) In addition, the PI presented a new proof of the convergence of Google searching algorithm although there are several proofs available already. See Section 2.3 of Appendix I. This analysis is used for new algorithms based on sociological analysis discussed in 6) below.

*This project is supported by Army Research Office Grant # W911NF-11-1-0322, August 23, 2011–Aug. 23, 2012

[†]Dept. of Mathematics, University of Georgia, Athens, GA 30602. Email Address: mjlai@math.uga.edu

[‡]Dept. of Sociology, University of Georgia, Athens, GA 30602. Email Address: mjlai@math.uga.edu. Email address: sodawn@uga.edu

- 3) The PI proposed a local update algorithm which is new. However, its computational efficiency is not compared with other updated algorithms yet. See Section 2.4 of Appendix I.
- 4) Certainly, the PI studied how to improve the efficiency of the computation in hoping to speed up the search. Due to the explosive increase of the webpages and the Internet surfers, such a study is definitely necessary. However, the study does not lead to any meaningful results so far.
- 5) The PI discussed with the co-PI about the web search algorithms in sociological sense which is one of the key points of this project. The PI found that the Bonacich centrality measure is the same as the one used in Google search algorithm if the relation matrix is the adjacency matrix with an appropriate normalization constant. See section 3 of Appendix I.
- 6) The co-PI suggested to use not only the out-linkages, but also the in-linkages in the search algorithm. In addition, the co-PI suggested to use the secondary linkages to further help determine the relevance in PageRank. Based on these ideas, the PI has proposed a new version of searching algorithm which uses both in- and out-linkages, and proved that the modified algorithm is still convergent and the convergence rate is determined by the damping factor. See section 4 of Appendix I.
- 7) The co-PI reviewed a lot of literature in sociology related to the PageRank, the relevant scores for linkages, in particular, for social linkages, as well as the social strong and weak ties, social cohesion, position and distance. See Appendix II.
- 8) The co-PI was willing to write it out a section to explain why one needs to use both in- and out-linkages in sociological terminology. However, the PI has not received such a write-up as Jan. 15, 2013.
- 9) Finally, the PI studied the Yahoo search algorithm based on its patent although the current Yahoo updates its search algorithm and the PI did not find enough time to study it. See Section 5 of Appendix I.

3 Conclusion:

The PI has surveyed all the mathematical literature related to the search engines up to today and the co-PI has summarized all the sociological concepts related to the internet search, not only the relevance among webpages, but also the closeness, betweenness among people, groups, etc.. The goal of this project to understand the PageRank sociologically and how it could be adjusted to be more effective sociologically as well as how to make the algorithms more efficient. This is still a central issue. The PI is not able to complete this goal of the project as the PI and co-PI have managed their time to only do the surveys of the current literatures in mathematical and sociological senses and have not found enough time to discuss their connections. They made a small progress towards the goal as mentioned above. The PI would like to continue this project if the ARO is willing to fund it again. Due to this project, the PI gains an excellent knowledge on the mathematical study of the Google searching algorithm and other search algorithms. He is ready to attack some research problems related to the algorithms such as how to speed up the computation and how to update in parallel the Google matrix, its PageRank vector and etc.. Also, he is now able to work with a sociologist better as he knows many sociological definitions and notations. See Items 5) and 6) above as an evidence.

With more time, this research team could more deeply digest the existing information, and propose some new algorithms, making use of sociological insights to improve mathematically the characterization of relations in large, complex systems. The PI and co-PI plan to continue their joint work toward a better understanding of the searching algorithms and discovering how to better situate them in the larger contexts of existing mathematical and sociological knowledge.

Appendix I

Ming-Jun Lai and Dawn Robinson, A Mathematical and Sociological Analysis of Google Search Algorithm, a manuscript, 2013.

Appendix II

Dawn Robinson, A Summary of Sociological Concepts Related to Social Network and Its Techniques for Quantifying Social Cohesion, Social Position, Social Distance, a manuscript, 2013.